

# White Paper Report

Report ID: 103966

Application Number: HD-51362-11

Project Director: David Pfeifer (depfeife@iupui.edu)

Institution: Indiana University, Indianapolis

Reporting Period: 9/1/2011-2/28/2014

Report Due: 5/31/2014

Date Submitted: 5/31/2014

**NEH Grant Number: HD-51362-11**

**WHITE PAPER**

**STEP**

**An Open-Source Scholarly Text-Editing Platform  
for Critical Editions**

**Project Director: David E. Pfeifer  
with André De Tienne and Young Chang**

**Indiana University, Indianapolis**

**31 May 2014**

## Building STEP, A Scholarly Text-Editing Platform for Critical Editions

### *Preamble: The horizon behind and beyond STEP*

The development of STEP, an online Scholarly Text-Editing Platform, was born of several needs experienced by the various editions in the Institute for American Thought (IAT) at IUPUI, most centrally by the Peirce Edition Project, whose publication of the Peirce Papers entails a level of technical and documentary complexity of uncommon magnitude.

Faced with the need to revamp their technological infrastructure from the ground up because of the disappearance of FrameMaker + SGML from the Mac platform, and with the realization that the solution had to be long-term, durable, and in phase with online and digital expectations becoming more pressing and sophisticated every day, the Peirce editors decided to launch the creation of a new production platform whose very conception would position the IAT editions (and other interested scholarly editions) well ahead of a curve we had all fallen behind.

What was also needed along these lines was a rethinking from the ground up of the coordination of our production *and* dissemination methods in the light of the kind of services our combined growing international constituency, scholarly or non-scholarly, had come to expect given recent and rapid technological advances. We needed to reassess some of our fundamental goals and reinvent or redesign the very products we contribute to scholarship so that they may be more easily integrated in a sophisticated digital environment suited for incoming generations.

We concluded that the successful scholarly editions of the future—successful both in terms of their ability to attract funding and meet and even exceed user expectations—would be those that take full advantage of advanced contents management frameworks so as to *stimulate transformative scholarship while opening it to broader audiences*. The re-conception of our work came therefore with the realization that we needed to view ourselves as permanent agents and partners in the advancement of knowledge. To do so successfully, we needed to design a *production* platform that would be in phase with what would ultimately be its natural target, a *dissemination* platform that could offer a wide and conveniently centralized array of options to users, including powerful search tools, interactive tools, collaborative tools, work zones, discussion areas, and feedback areas where user contributions get peer-reviewed, assessed, and professionally accredited.

Considerable work was accomplished over the last year by a team of HCI specialists at IUPUI to prototype that kind of dissemination platform, named CORPUS (for Collaborative Online Research Platform for Users of Scholarly editions). Although CORPUS is not discussed in this white paper since its conception was not part of the NEH grant, it is here mentioned only because STEP cannot be fully understood without keeping in mind that it has been designed *both* to help our editions produce their print editions as before, and also to produce digital editions reconceived and redesigned to suit CORPUS's goals.

This white paper discusses some of the challenges we encountered and the solutions we devised while developing STEP.

### *1. The challenge of hiring affordable competent people*

Designing and programming STEP involved many people over nearly three years full of ups and downs. If there is any lesson our grant experience can teach others (besides the technical accomplishments we managed to reach), it is that when it comes to hiring web developers and/or programmers in an academic environment, the principal hurdle is to find people who are very competent, who work hard, who understand and appreciate ambitious and unusually complex goals, and who can be counted on to stay and remain focused on task not just for months on end, but for a much longer haul. The principal obstacle in the academic world, especially in its sub-region called Humanities or Liberal Arts, is our incapacity to offer a competitive salary for a kind of position and responsibility that is generally very well compensated in the job market. As a result, we cannot expect that the persons we

will end up hiring will have all the competences we need, for those who do are not in academia. Another obstacle to be well aware of is the reality that the web developers/programmers we do hire will inevitably be called upon to execute a myriad other tasks that are entailed by their technological or computer-maintenance competence: they will be asked to help fix this server, that computer, that printer, to redesign a backup system, to be the point person on university technological policies, to work on this or that operation's website, etc. In other words, it is not good to be naïve about it: to hire a web developer/programmer is to hire someone who will end up working on your project only part-time, per force.

This is all to say that when we started the grant we had not foreseen that we would need to employ several consecutive developers, with an inevitable lapse of several months between each hire, plus the time needed to bring the new person up to speed in a constantly evolving project.

## *2. Learning Drupal and especially TEI*

It is one thing to keep recruiting personnel, and another to provide them with adequate continued training. All our developers were sent to attend many workshops in various places (including Oxford, UK) to improve their grasp of Drupal 7 on the one hand, and especially of TEI on the other hand. Scholarly editors discovered that in order to lead the technical team efficiently and provide them with solid direction, they had themselves to become involved very deeply in the design of every component of the platform—and, therefore, to become well versed, not so much in Drupal than, crucially, in TEI encoding. TEI encoding comes with a stiff learning curve. Understanding its logic is a scholarly editor's duty, and it is important (as another lesson to other projects) that edition leaders recognize that necessity. Programmers and web developers cannot be expected to appreciate all the intricacies of TEI. That is the responsibility of editors, who then use their understanding of it to help redirect or correct essential elements in the design of a platform meant to be TEI-compliant. One consequence is that the TEI P5 Guidelines became the favorite bedside reading of one of our edition directors, at least for a while, and not because they induced sleep—on the contrary, their logic was too enthralling for that.

We came to realize that STEP needed to be embedded with pedagogical features helping users (transcribers, textual editors) learn and understand what tags to use in what kind of textual settings, how to look for them, and how to remember their syntax. As a result, the TEI Editor includes ways of getting on-the-fly information about the TEI tagging system.

## *3. Improving the TEI XML Editor and Authoring Tool*

A key lesson learned from our working with HCI-aware students and developers was that the design of a platform as complicated and work-intensive as STEP had to take into account of all sorts of usability requirements and ergonomic conveniences. We therefore put STEP to the test, and asked our own transcribers to transcribe Peirce manuscripts in STEP in order to advise the technical team of what was working well and what was unpleasant or frustrating. Having long been accustomed to (and spoiled by) FrameMaker+SGML—which was after all a very well designed solution despite its complexity—our transcribers had high expectations. They could not really expect a brand-new browser-based online platform to behave all at once with the speed and reliability of a proprietary Adobe product. But they could certainly indicate what features they could not do without, and what ought to be their location within the interface so that they be easily accessible without the grievance of extra clicking or mousing. Their comments proved a great help in that regard. As a result, our current TEI XML authoring tool simplifies the creation process through context-based editing. Its minimal user interface aims to provide only those tagging options (elements and attributes) that are validly available at any current point of inscription within a document based upon its underlying structure and XML schema. Redundant methods of input have been built in to accommodate varying user preferences.

#### 4. *Improving the editorial workflow*

A customizable editorial workflow has been configured to support common critical edition activities, from transcription to editing to annotations. STEP supports both individual and collaborative work on TEI documents while preventing accidental overwrites. Critical editions are able to modify the workflow to support specific needs. An important lesson we gained here is that an online workflow within a CMS endowed with versioning and document routing provides advantages over a paper-based workflow that are such that it would be a mistake to try simply to replicate our paper-based workflow (inherited from decades of practice) into that online workflow. As a result—and this is still ongoing—we have been forced to fundamentally rethink our workflow. This means re-envisioning how documents are worked on, how proofreaders do their work, how corrections are entered, how many times, and how all of that is organized and structured within the CMS in ways that make sense for online users (ways that on top of that need to be customizable to suit the requirements of different editorial practices). We have so far redefined our transcription workflow in that regard, but have yet to redefine other workflows such as that entailed by critical editing.

#### 5. *Versioning, revision, document routing, and access control*

STEP has been set up to save separately each iteration of a document whenever the user clicks the Save button, regardless of its routing stage. Any version may be viewed and reverted back to a previous state by authorized staff. Several categories of users with different levels of read/write access can be preconfigured within STEP in conjunction with the different editorial workflows, and all such features are customizable. Each iteration of a document will come with a routing status that exhibits what stage the document is at within the workflow, who are the users who have it currently on their workbench, and which user or set of users the document will be routed to next.

#### 6. *Highlighting and comments*

A digitized form of highlighting sections or regions of a text or image was implemented. The types of highlights and the colors associated with them are customizable, depending on editors' needs. They provide ways of annotating textual passages or image segments to indicate different classes of editorial interventions done or yet to be done, or to draw attention to anything in particular. A few minor programming bugs in this area between the user interface and the back-end program have been worked out. Subsequent (and very recent) reflection about these highlighting and commenting features made us wonder whether the niftiness attached to this technique of highlighting textual passages to attach to them such things as alteration descriptions, emendations, textual notes, and annotations, was actually desirable. Our growing suspicion is that, even though it has become a fashionable method, it may actually not be recommendable at all because the technique *appears* (pending further inquiry) completely to eschew TEI XML tagging. What the technique creates are very long Jason-based XML strings that have no resemblance to TEI tags. If we are right about it, then this aspect of the TEI editor will need to be revamped (more about in subsection 9 below regarding “STEP Tools,” a suite of add-ons to STEP). We may decide to leave the system in place as an option for other editions, while also offering a more rigorous TEI-compliant system.

#### 7. *TEI XML validation*

There are two types of XML validation: well-formed XML and schema-valid XML. STEP's TEI-compliant authoring tool minimizes the risk of creating either non-well-formed XML or invalid schema XML thanks to its context-sensitive editing features and its spare user interface. It also provides immediate visual alerts when non-well-formed XML is manually entered in the editing window. Back-end programming has been completed to process complete document schema validation prior to saving. What remains to be done is testing the limits of TEI XML validation. We have noticed that when very complex sets of TEI tags (involving multiple levels of nested alterations) go through this sieve,

validation chokes even though we know that the logic governing these complex nests of tags is irrefragably correct.

#### 8. *Upload and management of high-resolution digital images*

Digital images of original manuscripts and any supporting images, charts, figures, tables, and other graphics that may be published as part of a given text can be uploaded into the system through the STEP interface. Manuscript images can be referenced, zoomed in and out, and moved around in the web-browser viewport during document creation. They can also be highlighted and have comments attached to them much as if editors had physical copies on their desk, though the highlighted sections and comments are not digitally marked on the image itself to preserve its integrity for subsequent viewings. These highlights and comments can then be viewed and referenced by subsequent editors throughout the editing process.

#### 9. *STEP Tools — an essential improvement to our STEP development strategy*

STEP was initially conceived to be a comprehensive and self-sufficient scholarly editing platform. Over time it became evident that STEP could be comprehensive *but not self-sufficient*. STEP development followed the principles and protocols of Human-Computer Interaction (HCI) design, including a series of usability tests. Once the online TEI-XML editor was developed, testing showed that it was one thing to use the TEI editor to transcribe relatively simple texts with few authorial corrections. But when it came to complex copy-texts strewn with multiple levels of insertions and deletions—as is the rule with Peirce’s manuscripts—the difficulty and tedium of tagging complex alterations became quickly an unbearable burden on transcribers, professional or not. TEI tagging is not for the faint of heart, if it is to be done as correctly as possible—and exact TEI tagging is what STEP is about.

And so arose a pressing need for specialized tools that would alleviate the tedium, automatize complex tagging procedures, and decrease errors. It became evident that transcribers could not be expected to master all aspects of TEI nor to enjoy creating a forest of tags within documents while maintaining their sanity. The need to retain them and to flatten a stiff learning curve required, therefore, that tagging tools be programmed in such a way that users would be spared the labor of tagging textual elements (including attributes and values) and anchoring tags to the largest degree possible.

It also became clear that programming such specialized tools would be arduous. Take for example a specialized TEI-XML transcription tool. If well conceived, such a tool should enable users to overcome many difficulties and achieve many goals. Here is a sample list:

- (a) Have ready at hand, at the fingertips, several options for entering TEI tags and attributes: pull-down menus, pop-up menus, clickable buttons, autocomplete features, self-correcting features, keyboard shortcuts, programmable function keys, efficient connection to relevant online TEI documentation, and access to examples.
- (b) The flexibility to import the relevant XML document from the STEP platform into the specialized tool, but also the option to import any other untagged transcription from any other source.
- (c) Ways of tagging multilevel alterations (say, a phrase deleted and replaced five times through superimposed deletions and insertions, partial or full) through tag nesting without having to type a single angle bracket.
- (d) Ways of tagging complex transpositions, including their special TEI attributes and anchors.
- (e) Automatic tagging and anchoring of deletion, addition, and damage spans.
- (f) Easy ways of inserting special characters with diacritical marks, foreign characters, symbols, unusual punctuation, fractions, in automatic HTML code.
- (g) Automatic administration of all TEI-approved anchoring systems, including the correct syntax and placement of abbreviations and numbers.
- (h) Automatic generation of *syntactical* descriptions of alterations based on their tags, attributes, and anchors, in good English.

- (i) Automatic generation of *pragmatic* descriptions of alterations, ready for the critical apparatus, also based on an intelligent translation of tags and attributes.
- (j) Ability to change and customize all aspects and features of the tagging tools, including the assignment of colors to distinct classes of tags, the arrangement of tags within pull-down menus and pop-up menus, the addition or deletion of tags, attributes, and values within those menus.
- (k) Ability to export any content of any field associated with the tool either to the STEP platform or to any other text or browser software.
- (l) Ability to view the tagged transcription in run-on or indented view, with or without the anchors, and with or without the tags.
- (m) Ability to completely separate the tagging of authorial alterations from the tagging of descriptive elements in order to achieve tag clarity while avoiding the scourge of overlapping tags (an XML restriction).
- (n) Ability to consult a digitized image of the source text on screen while transcribing.

The STEP platform implements several of those requisites partially or fully, but not all: requisites requiring sophisticated programming (such as c, d, e, g, h, j, l, m) are not within its reach. Two thoughts occurred to us: (1) that all of those requisites could be programmed externally far more efficiently than within the browser interface; and (2) that it was actually perfectly acceptable for some STEP tasks to be accomplished outside the platform if we could establish a good and safe protocol of two-way data transfer between the platform and the external tool. From the standpoint of transcribers and users, it is really immaterial whether the software they use is an online browser, an app within their workstation, or a combination of the two.

What mattered, therefore, was to find a way of programming those external tools (or apps) fast, efficiently, using a web-aware, open-source, community-supported, cross-platform programming environment compatible with online browsers, capable of communicating with other applications, of launching them, of opening, reading and writing processes, and of executing shell commands. Several options were considered but in the end we chose one that was most appealing for its power and reputation for rapid development: LiveCode, a software developed by Runtime Revolution (aka RunRev), a company based in Edinburgh, Scotland. RunRev was founded in 1997 with the goal of enabling fast creation of applications for enterprise, commercial, creative and academic environments. LiveCode, that company's cross-platform programming environment, was launched in 1998, initially under the name Revolution. It began as an expert Integrated Development Environment (IDE) for a graphical user interface and development environment that supported UNIX, Windows, and Mac. Over time, the technology developed into software that could be deployed to mobile platforms (iOS and Android) as well as desktop (Mac, Windows, Linux). In 2009 a server management system was added and Revolution was renamed LiveCode in 2010. In 2013 RunRev launched an ambitious initiative to provide an open source version of the LiveCode software. The first version of this, [LiveCode Community](#), launched in April 2013. The LiveCode open source edition includes the entire code base for all platforms, including the development environment and documentation.

The beauty of LiveCode comes from its fourth-generation scripting language, which is English-like and therefore highly readable and easy to debug and correct. LiveCode allows programmers to develop and edit a program while that program is running (no wait time, no separate compilation). The consequence is that software created in LiveCode can be easily shared with everyone, and understanding what another programmer has done in that language is not complicated: reading the program in English is all it takes (almost). This extends the open-source concept from the specialists to the non-specialists.

Having adopted LiveCode as our open-source solution, we began developing "STEP Tools," a suite of STEP add-ons designed to augment the platform with specialized TEI tagging systems. A STEP tool is not a plugin but a full-blown external application. Although the initial idea was to simply prototype each tool so that it could be programmed later as part of STEP, the first prototype instantly turned into something functional precisely because it was done in LiveCode. From that point on, the applica-

tion took a life of its own, and in three months “STEP Transcripator” was born, the first of a suite of tools. *The PDF mentioned at the end of this white paper provides a comprehensive presentation of STEP Tools.*

We have at the moment four STEP Tools either done or in the works.

(1) **STEP Transcripator** is by far the most complex; it is now fully functional and almost completely done (only one of its sub-tools remains to be programmed: a tool that automatically tags TEI-defined “alternations,” which entail the most convoluted kind of TEI transcription tagging there is). *A complete user guide to STEP Transcripator with hundreds of screenshots is found in the PDF just referred to.* It is designed to improve transcriptions begun in STEP. An Export button in STEP’s transcription window allows STEP users to export the core content of the TEI transcription (the one between the `<body></body>` tags) into a file with extension “.step”. That file can be dropped on the icon of STEP Tools’s application, which will launch STEP Transcripator with the content of that file in the transcription field, ready for advanced tagging. Once a text has been fully tagged, an Export button in STEP Transcripator turns it into a new .step file. That file can in turn be imported back into STEP thanks to the Import button also found in its transcription window. STEP Transcripator can also be used independently of STEP by anyone in need of transcribing any text in TEI-compliant fashion. Plain-text, RTF, Word, OpenOffice, or HTML files can be moved into STEP Transcripator, and the latter will output its tagged transcriptions in plain-text, RTF, HTML, or XHTML form (but all really being XML documents, although other options are available). A new text or transcription can be started from scratch in STEP Transcripator as well. What STEP Transcripator does is *everything that has been listed above, from (a) to (n)*. It is a comprehensive TEI-XML compatible solution completely specialized in and focused on the act of transcribing.

(2) **STEP Emendator** has so far been prototyped, and *the prototype is presented in the PDF mentioned above*. STEP Emendator is a tool to help textual editors create and tag emendations and textual notes using several collated sources of authoritative texts after transcriptions have been perfected. Its goal is again to allow textual editors to do their work without having (ideally) to type a single angle bracket. For instance, the idea is to have the textual editor highlight a word in need of editing, type the correction in a special field, and click a button that will automatically create the emendation entry within the copy-text (using `<corr>` tags), while also generating an apparatus list of emendations (using `<app>` tags) that takes into account the type of emendation each one is (e.g., spelling or punctuation corrections). STEP Emendator will provide the correct tags and attributes, as well as the full complement of lemmas, square brackets, sigla, glosses, and anchors necessary for cross-linking the emendations lists in the critical apparatus with the author’s text. Again, much of this can be done directly in STEP, but the tool is intended to make the life of textual editors far easier.

(3) **STEP Image Browser**, a tool that can be used along with any of the other tools (it can be made to float above them); its function is to provide visual access to digitized images of manuscripts, for transcription, critical editing, or annotation purposes. STEP Image Browser is also fully functional at this writing, although other features like a zoom will be added to it in time. *It is presented in the PDF as well.*

(4) **STEP Text Comparator**, a tool that can be used along with STEP Emendator to the extent that emendations rest in part on the collation of variant texts and the adoption or rejection of substantive variations in post-copy-texts. The text comparator helps visualize differences between two texts at a time, in either direction, on the basis of three distinct ranges of comparison: by word span (from 2 to 8 words at a time), by sentence segments, and by full sentences. *It, too, is presented in the PDF.*

Other STEP Tools are under consideration, chief of which is **STEP Annotator**. We will turn to it when STEP Emendator is done. STEP Annotator will be a tool specialized in the administration, collection, and tagging of scholarly annotations to an author’s text (data collection, drafting of editorial scholarly notes, selection of notes, editing, correction, layout). STEP has a module in place for it though not functional yet. Based on our experience with many graduate students and more established scholars who have been drafting annotations for the IAT editions over several decades, it is already



clear, however, that we will also need a truly dedicated annotation-tagging system that reduces the tedium of complex tagging to a minimum (which is directly correlated with a sizable reduction in the amount of errors). The tool needs to be connected to archival sources, to collections of annotations in previous volumes of an edition, to indexes of previous volumes, etc. The range of TEI tags it will need to give access to is vast, and a system will be needed to administer that range according to topical categories.

### *Accomplishments*

Six technical goals were outlined in our initial grant proposal. We review each one of them below, and gauge what we have accomplished.

(a) *Importation of digitized images of original documents.* STEP provides a widget that imports images in png, jpg, jpeg, and gif formats and makes them selectable through thumbnails. Clicking a thumbnail brings the corresponding image into a viewing area facing directly the transcription field. Tools that slide in and out of view allow users to mark up image areas; zoom in and out; pan up, down, left and right, and resize images. Whether images are from original documents or from copies of them depends only on what holding archives have digitized and opened to public in that form. STEP Tools also provides an image browser and viewer that allows importing images of a wider range of formats; the import is automatic as soon as an image folder is selected. It does not offer yet image manipulation tools. *Goal (a) has therefore been accomplished (100%).*

(b) *Production of rigorously exact transcriptions.* Exactitude of transcriptions means two distinct things: exact reproduction of the source text, and exact TEI tagging of that text. Exact reproduction depends in great part on a transcriber's deciphering and typing skills, which are beyond the scope of programming (no spelling checker is allowed, for we need to preserve an author's spelling errors in transcriptions). But it is aided by a well-conceived work area that is not creating unnecessary distractions or sources of confusion, and we think STEP has achieved that. As to the rigor and exactitude of TEI tagging, STEP provides easy access to tags, autocompletion of tags, TEI XML validation, and access to TEI information regarding tags. In addition, STEP Transcriptor is a dedicated add-on that specializes in TEI tagging of alterations and also in the application of descriptive TEI tags (see the appendix for a fundamental distinction between transcriptive tags and descriptive tags). STEP Transcriptor also autocompletes tags, but does a great deal more as well, including the automatic production of lists of alterations ready for the apparatus. *Goal (b) has therefore been accomplished (100%).*

(c) *Speeding up of collations and manipulations of texts.* There is no full-blown STEP Collator drawn up yet on our blackboard. But we have developed STEP Text Comparator, which offers ways of visualizing differences among texts that go beyond text comparators available online (all of them based on a partial grasp of the *diff* shell command; STEP Text Comparator does not rely on *diff*). STEP Text Comparator automatically generates lists of differences in which the lemmas come from the base text, and the glosses represent the readings from the variant texts. Let's say that *goal (c) has therefore been partly accomplished (40%)*. A true "historical collator" application would offer a great deal more, including a file management system and specialized TEI tagging. Developing at the moment is not a priority however.

(d) *Enabling the online scholarly editing, annotating, and formatting of texts in a WYSIWYG interface that keeps track of and archives every iteration of a document through multiple stages of corrections and editorial interventions.* Thanks to the Drupal framework that sustains STEP, the workflow system that date-stamps and keeps track of each iteration of a document through any particular stage of production is in place. It is at the moment functional as far as the production of text transcriptions goes. But it is not yet functional as far as the critical editing of texts goes. The editing module is formally in place but not developed to the same point yet because we have focused our attention on getting the first module (transcriptions) as right as possible, reasonably thinking that once we have worked the kinks out of it, it will only be a matter of duplicating it, give and take some customization, into other modules. *Hence, goal (d) has been only partially achieved in the platform. Combined with*

*our advanced prototyping of STEP Emendator, however, it would be fair to gauge our accomplishment in this crucial area at 50%.*

(e) *Linking edited texts and their components to the digitized documents on the one hand, and to their critical editorial apparatus on the other.* This is a technically complex undertaking. A highlighting system (described in the first part of the PDF mentioned earlier) has been developed that links highlighted portions of a transcription to notes provided by the transcriber—notes that can be anything: descriptions of alterations, emendations, scholarly annotations, special cases of hyphenations, and the like. STEP Tools, on the other hand, and at this stage STEP Transcriptor only, takes into account all the TEI rules regarding cross-linking and anchoring apparatus list items with the source text relative to the desired output (print or online). STEP Transcriptor has been programmed to generate and update those complex anchors (complex due to variable attribute and value syntax) automatically. We have therefore acquired a solid understanding of how the cross-linking protocol works, and developed algorithms that take care of it. It will remain to apply this knowledge to both the critical editing module and the annotations module. We do not anticipate this process to be overly difficult since we have learned how to deal with it. *Hence we gauge this goal to have been accomplished at 70%.*

(f) *Streamlining the conversion of copy-texts to fully laid-out and hyperlinked texts readied for online or printed publication.* This is STEP's ultimate holy grail—not so holy because it is of course reachable, and everything we have done goes in that direction. But we are not there yet. *Hence this goal can only be said to have been plowed on toward by 20%.*

To sum up, on the one hand we have not accomplished to the full all of our goals, but on the other hand we have managed to accomplish a great deal more than announced. This is due to the fact that in time we realized we had to revise our initial strategy drastically owing to the lesson learned from our usability tests. That lesson was that it was not sufficient to implement an online TEI-XML editor (something which by itself is quite a technical accomplishment). To be at all usable and workable, that editor needed powerful enhancements so as to make the work of transcribers and editors bearable. Our answer is STEP Tools as a suite of add-ons. That was no part of our initial grant proposal. But as demonstrated in our PDF documentation, those tools do pack a great deal of punch.

Apart from all of this, other challenges remain on our workbench.

In the first place we need to finalize STEP's front-end interface design. Most of the work to date has been in the development of the back-end structure and functions of the TEI XML editor and of the Drupal content and workflow management features. We are acutely aware of the fact that the interface itself needs to be functionally and esthetically improved, and that is one of our next priorities.

In the second place, we are continuing to rethink the best ways of managing the major challenge of transforming the digitally ready documents into print book form. The complexities involved in such a transformation (passing from STEP to layout and printer-ready copy) are still under examination, with a solution in progress. One specific challenge is to allow all editors to fulfill their specific tasks while tagging the same underlying text (which itself changes over time as it gets emended and edited) without interfering with one another's work. The solution now decided upon (following the creation of STEP Transcriptor and the realization that we needed to sharply separate transcriptive from descriptive tagging) is to create the equivalent of what is called a layer in image-processing software. Each module adds a distinct layer of tags to the same underlying text, and that layer is kept separate from the other layers. The result to be achieved is twofold. In the online digital edition, users will have access to each layer (alterations, emendations, rejected substantives, etc.) separately without confusion and without even being aware that they are looking at a different document since the underlying text remains identical. In the print edition, on the other hand, the textual apparatus for each selection will pull its distinct lists of selected alterations, emendations, textual notes, annotations, etc., from the distinct layers as well, again without risk of confusion. This separate layering of types of apparatus tags will therefore help decrease errors in the creation process by simplifying the tagging of complex documents. It will also allow for asynchronous document creation among editors, since the creation of dis-

tinct apparatus lists (say, philosophical annotations) will not need to wait for other lists to be completed (say, emendations).

In the third place, the various components of STEP will require some configuration prior to implementation. Development of a downloadable and distributable application that facilitates basic installation while providing various configuration options is underway (for both Mac and Windows). For any edition that wants to customize STEP beyond the out-of-the-box experience, special documentation and extra assistance may remain necessary to help implement STEP to full user satisfaction.

### *For whom would STEP be useful?*

Our technical editor has been attending several TEI workshops during the winter and spring terms, and he reports that whenever he describes to workshop attendees what STEP or STEP Tools can do, everyone, including the specialists, are deeply impressed because even though a great deal has been done over the last 20 years within the TEI organization and its affiliated projects, no one has yet produced tools as specialized as ours. We are aware that others are groping in that direction, but also aware that we are ahead.

From the start we have been developing STEP, and also CORPUS, in a spirit of generalization. Those tools are not confined to our own local needs. Granted, if we manage to transcribe and edit Peirce's diverse philosophical and scientific manuscripts using STEP, then it follows that STEP can be useful to a wide range of editions right from the start. But we always keep in mind that every edition is different, and that nearly every feature we implement needs to be customizable. A good case in point is STEP Transcriptor: Appendix 2 contains several sections that show how to customize every aspect that matters in that particular application. Our applications are thus designed to be serviceable beyond the IAT editions, and therefore we believe that part of our audience consists of many other scholarly editions, at least those that are contemplating the digital transition, which is pretty much all of them as is clear when attending the meetings of the Association for Documentary Editing (ADE).

The larger audience is the one that will benefit from STEP's ultimate goal, which is simply to help produce solid scholarly editions on paper and online. That audience will not care about the means but about the scholarly texts produced by the means. It is for them that STEP, and CORPUS, exist.

### *Continuing Plans*

We certainly do plan continuing the project simply because STEP is the technological future of our editions, and especially of the Peirce Edition Project, a scholarly edition with a mission that will take a few more decades to accomplish. STEP is a strategic need of the IAT, and thus also of the School of Liberal Arts at IUPUI.

Indiana University as a whole, and IUPUI in particular, have been developing institutional mechanisms to support initiatives in the digital humanities. IUPUI created for instance [IAHI](#), IUPUI Arts & Humanities Institute, which offers internal grants. We applied for and received one of those grants two years ago, which enabled us to collaborate with the School of Informatics to develop the CORPUS prototype successfully. We intend to apply again to begin part of the programming of CORPUS.

We have also begun collaborating with scholarly initiatives at Indiana University-Bloomington, including the [Catapult Center for Digital Humanities & Computational Analysis](#) (where we presented CORPUS last year to their great interest) and also with IU's Institute for Digital Arts & Humanities ([IDAH](#)).

We have mentioned above our intention to pursue an NEH Digital Humanities Implementation Grant. Another complementary possibility is to apply at some point to the Mellon Foundation's program called "Scholarly Communications and Information Technology." We think that the combination of STEP and CORPUS, once it has reached a more advanced stage than is currently the case, is likely to become a model of interest to that program, whose goals are congruent with what we are seeking to design. Indeed, CORPUS's central goal is to provide end-users with a set of tools that will stimulate

scholarly research. Several of those tools (some of which exist already elsewhere) will allow the navigation of our critically edited texts across different layers of authorial drafts and variants, while also bringing on screen images of original manuscripts and annotations linked to discrete textual elements. Other tools will allow users to provide comments sharable with different readership segments to further annotate texts, supply scholarly commentary, suggest corrections, discuss the editing, take notes, customize the configuration of texts for personal convenience, communicate via a specialized forum, propose links between text segments and online literature, or collaborate on specific tasks.

### ***STEP's Positive Impact***

The major impact in the long term is providing the tools for many editions to transition smoothly from print to digital production—smoothly, because STEP should make both possible: one is not opposed to the other. We have learned they are actually complementary, serving different purposes and adjusted to different modes of work and research. Both STEP and CORPUS have become well noted at IUPUI as a model example of digital humanities initiative inspiring active collaborations among schools. The faculty and the students who have been collaborating with us have been enthused by the unique challenges of our project: they see it as setting a high standard, well fitted for the kind of rigorous and creative training they seek to give or receive. Students have been thankful for the opportunity STEP or CORPUS have provided them, and, partly as a result of it, all of them have been able to find employment in notable institutions. The university hierarchy is well aware of this impact, which also played a role in the IAT receiving permanent “signature center” status.

### ***Final Product***

STEP-beta (the platform + STEP Tools) will be made eventually available to other interested scholarly editions. Regarding the dissemination of the ultimate product, when the dissemination platform (CORPUS) gets finally completed, or at least in a satisfactorily working condition, it will be combined with STEP into an integrated solution and again be made available to other interested scholarly editions. Should that happen, the usefulness of that integrated solution to the field should be considerable and have noticeable impact on the evolution of scholarly methodologies and communication in the humanities.

David E. Pfeifer, Grant Director

André De Tienne, Director and General Editor, Peirce Edition Project

Young Chang, Technical Editor, Institute for American Thought

Indiana University, Indianapolis, 31 May 2014

**The Institute for American Thought will gladly send a PDF combining two documents: (1) STEP: Technical Guide, and (2) STEP Tools: A Suite of Add-ons to STEP. Please send your request by email to [iat@iupui.edu](mailto:iat@iupui.edu)**